COMMITTEE T1 - TELECOMMUNICATIONS STANDARDS CONTRIBUTION

DOCUMENT NUMBER:	T1A1.5/97-102 T1A1.7/97-003
T1BBS FILES:	7a151020.doc, 7a170030.doc 7a151020.ps, 7a170030.ps 7a151020.pdf, 7a170030.pdf
DATE:	January 16, 1997
STANDARDS PROJECT:	T1Q1-12, T1Y1-20
TITLE:	Additional information on proposed objective quality measure for the audio portion of an audio-visual session
SOURCE:	Institute for Telecommunication Sciences National Telecommunications and Information Administration U. S. Department of Commerce
AUTHORS:	D.J. Atkinson and Stephen Voran
CONTACT:	D.J. Atkinson NTIA/ITS.N3 Tel +1 303 497 5281 325 Broadway Fax +1 303 497 5323 Boulder, CO 80303-3328 email dj@its.bldrdoc.gov
DISTRIBUTION:	T1A1.5, T1A1.7, T1BBS

ABSTRACT: This contribution provides information requested by members of T1A1 pertaining to the performance of the MNB-based measure and the performance of the P.861 algorithm under a variety of network conditions.

Introduction

This contribution responds to the requests for further details on the objective quality measures proposed for the audio portion of an audio-visual session in T1A1.5/96-123 / T1A1.7/96-037[1]. Specifically, this contribution provides information on the performance of the MNB algorithms, in the form of three scatter plots that give further insight into the speech coding and channel conditions that comprise Test 3, as defined in [1].

Scatter Plots for Test 3

The scatter plots that follow show the relationship between MOS values and three objective measures. The output of the new ITU tool for the objective quality measurement of speech codecs, P.861[2] is shown in Figure 1. The output of MNB Structure 1 is shown in Figure 2 and the MNB Structure 2 output is shown in Figure 3. The key to the symbols used in those figures is given in Figure 4. Note that the per-condition correlation values agree with those shown in columns 2, 3 and 4 of Table 3 in [1]. Note also that Test 3 contains non-waveform coders and error conditions (randomly distributed bit errors), both of which are outside of the scope of applicability stated in Recommendation P.861.

P.861 generates Noise Disturbance values, which range from 0 to +∞ and the MNB structures generate Auditory Distance values which also range from 0 to +∞. These values can then be transformed into a finite interval using a logistic function:

$$y = \frac{1}{1+e^{a \cdot x + b}} ,$$

where x is the input to the logistic function (Noise Disturbance or Auditory Distance), and y is the resulting output. Note that this function has asymptotes at 0 and 1. In the case of P.861, the values of a and b were chosen to maximize the coefficient of correlation across all 7 tests described in [1]. The maximizing values are a = 0.5431 and b = -1.6716. Without this logistic function, the magnitude of the coefficient of correlation between Noise Disturbance and MOS values is reduced from 0.795 to 0.789. For both MNB structures, the values of a and b were chosen to maximize the coefficient of correlation for Test 1 and Test 2 only, as described in [1] (i.e., the MNB algorithm and the logistic function were "trained" on Tests 1 and 2). For MNB-1, the maximizing values are a = 1.0 and b = -2.4991. Without this logistic function, the magnitude of the coefficient of correlation between Auditory Distance and MOS values is reduced from 0.951 to 0.945. For MNB-2, the maximizing values are a = 1.0 and b = -2.2404. Without this logistic function, the magnitude of the coefficient of correlation between Auditory Distance and MOS values is reduced from 0.929 to 0.917. (The logistic function is not part of the proposed objective quality measures.) From the graphs, one can see that the majority of the increase in correlation achieved by using the MNB algorithms came from both errored and clear channel conditions with IMBE, STC and LPC coding algorithms, as well as those errored conditions with ADPCM, CVSD, VSELP and CELP coders.

Figure 1: P.861 Measurements for Test 3, Rho=.795

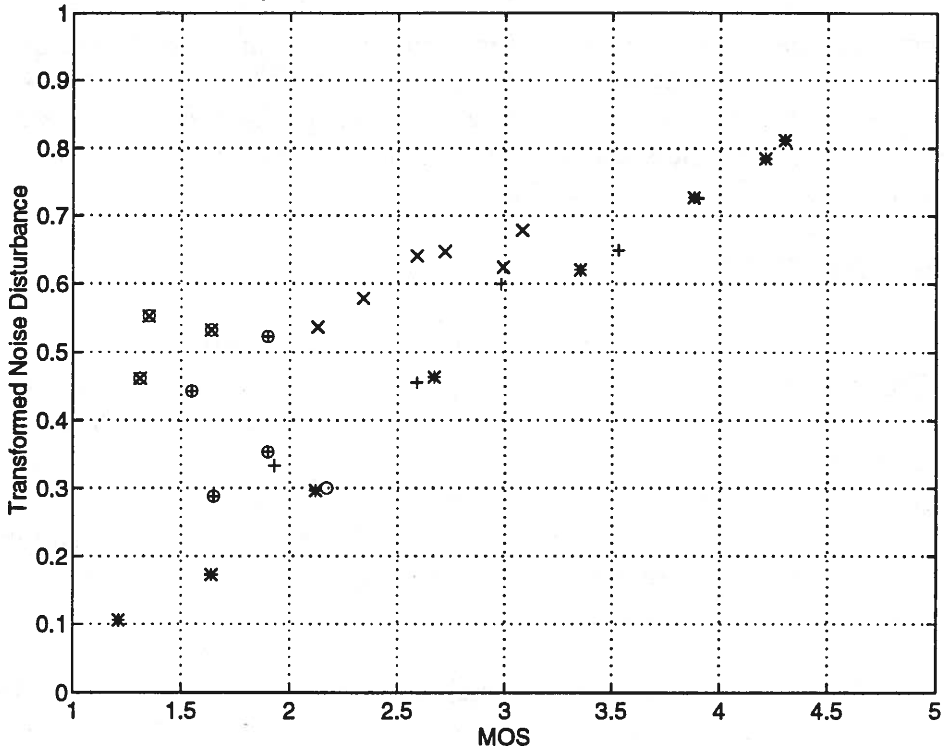

Figure 2: MNB Structure 1 Measurements for Test 3, Rho=.951

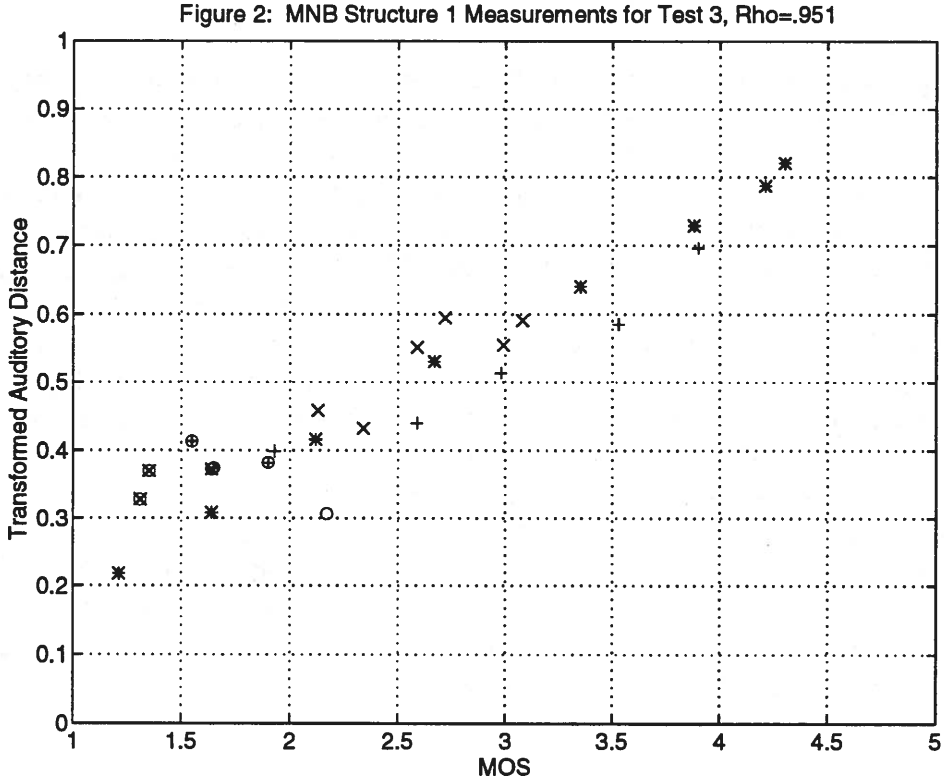

Figure 3: MNB Structure 2 Measurements for Test 3, Rho=.929

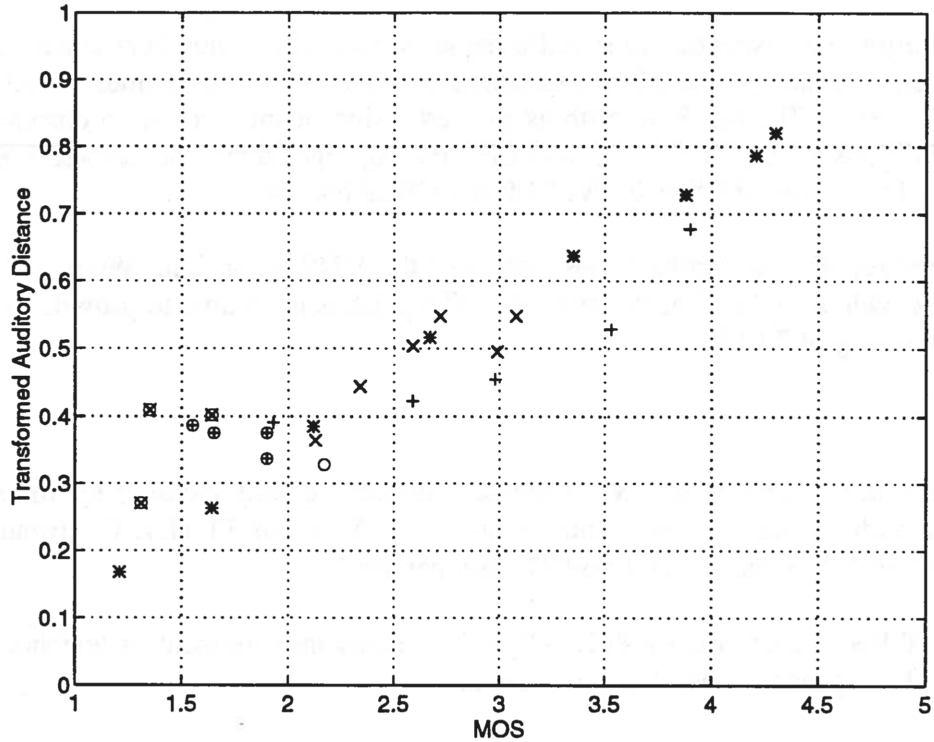

Figure 4: Symbol Key for Figures 1, 2, and 3

✱	MNRU
○	POTS
+	ADPCM, CVSD, VSELP, CELP
×	IMBE, STC, LPC
⊕	ADPCM, CVSD, VSELP, CELP, with errors
⊠	IMBE, STC, LPC, with errors

Summary

This contribution has responded to specific requests from T1A1 members concerning details of the performance results of the MNB algorithms and P.861 algorithm under conditions outside the scope of P.861. The MNB algorithms showed a significant increase in correlation for clear channel conditions with IMBE, STC and LPC coding algorithms and errored conditions with IMBE, STC, LPC, ADPCM, CVSD, VSELP and CELP coders.

T1A1 also requested an algorithmic description of the MNB algorithm. We regret that we were not able to provide it to T1A1 at this meeting. ITS expects to be able to provide the information at the next meeting of T1A1.

References

[1] Atkinson, D.J. and Voran, S., "Proposed objective quality measure for the audio portion of an audio-visual session," Contribution to T1A1.5 and T1A1.7, Contribution Numbers T1A1.5/96-123 and T1A1.7/96-037, October 1996.

[2] ITU-T Recommendation P.861, "Objective quality measurement of telephone-band (300-3400 Hz) speech codecs," Geneva, 1996.